

Attorney Docket No. 1359.1052

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Satoru WATANABE, et al.

Application No.:

Group Art Unit:

Filed: August 29, 2001

Examiner:

For: VOICE INTERACTIVE SYSTEM AND VOICE INTERACTIVE METHOD

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN  
APPLICATION IN ACCORDANCE  
WITH THE REQUIREMENTS OF 37 C.F.R. §1.55**

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

In accordance with the provisions of 37 C.F.R. §1.55, the applicant(s) submit(s) herewith  
a certified copy of the following foreign application:

Japanese Patent Application No. 2000-400480

Filed: December 28, 2000

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing  
date(s) as evidenced by the certified papers attached hereto, in accordance with the  
requirements of 35 U.S.C. §119.

Respectfully submitted,

STAAS & HALSEY LLP

Date: August 29, 2001

By: *H. J. Staas*  
H. J. Staas  
Registration No. 22,010

700 11th Street, N.W., Ste. 500  
Washington, D.C. 20001  
(202) 434-1500



日 本 国 特 許 庁  
JAPAN PATENT OFFICE

JC997 U.S. PTO  
09/940522  
08/29/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2000年12月28日

出 願 番 号

Application Number:

特願2000-400480

出 願 人

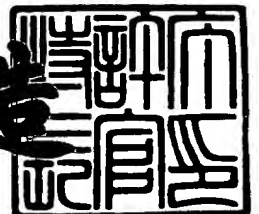
Applicant(s):

富士通株式会社

2001年 6月14日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3055811

【書類名】 特許願

【整理番号】 0095342

【提出日】 平成12年12月28日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 3/16

【発明の名称】 音声対話システム及び音声対話方法

【請求項の数】 10

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 渡辺 理

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 伊藤 映

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 紀伊 隆弘

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 木島 裕二

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100095555

【弁理士】

【氏名又は名称】 池内 寛幸

【電話番号】 06-6361-9334

【手数料の表示】

【予納台帳番号】 012162

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9803089

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声対話システム及び音声対話方法

【特許請求の範囲】

【請求項 1】 利用者端末から利用者の音声情報を入力する音声情報入力部と

前記音声情報について音声認識を行い、前記音声情報の内容を解析する音声認識部と、

前記音声情報の内容に応じて前記音声情報の伝達経路を制御する音声情報仲介部と、

知識データベースを参照して前記音声情報に対応した応答内容を抽出して、前記応答内容に応じた合成音声を作成する対話エンジンと、

前記合成音声を出力する音声情報出力部とを含む音声対話システムであって、  
前記音声情報仲介部は、前記利用者の対話が円滑に行われているか否かを常時監視し、前記利用者の対話が円滑に行われていないと判断した場合には、前記利用者と前記対話エンジンとの対話への他の端末からの第三者の参加を許可することを特徴とする音声対話システム。

【請求項 2】 前記音声情報仲介部は、前記利用者の対話が円滑に行われているか否かを、前記知識データベースに前記応答内容が見つからなかった回数の累積値が設定限界回数を超えたか否かによって判断し、前記設定限界回数を超えた場合には、前記利用者と前記対話エンジンとの対話への他の端末からの第三者の参加を許可する請求項 1 記載の音声対話システム。

【請求項 3】 前記音声情報仲介部において、前記利用者の対話が円滑に行われているか否かを、前記対話エンジンが応答してから前記利用者が反応を返すまでの平均反応時間によって判断し、前記平均反応時間が第 1 のしきい値を超えた場合、あるいは第 2 のしきい値より小さい場合には、前記利用者と前記対話エンジンとの対話への他の端末からの第三者の参加を許可する請求項 1 記載の音声対話システム。

【請求項 4】 前記音声情報仲介部において、前記利用者の対話が円滑に行われているか否かを、前記利用者の入力した前記音声情報の平均的な音質の評価値

によって判断し、前記平均的な音質の評価値が第 1 のしきい値を超えた場合、あるいは第 2 のしきい値より小さい場合には、前記利用者と前記対話エンジンとの対話への他の端末からの第三者の参加を許可する請求項 1 記載の音声対話システム。

【請求項 5】 前記音声情報仲介部において、前記利用者が対話を開始してからの対話時間及び対話エンジンとのアクセス度数に応じて対話進捗度を定め、前記対話進捗度が低い順に前記第三者の参加形態が、前記第三者へ前記利用者との対話内容を表示して前記第三者によって前記対話内容を更新する介入、前記第三者が前記利用者と並行して入力する並行入力、前記第三者が直接前記利用者との対話する切替と、順次変化する請求項 1 記載の音声対話システム。

【請求項 6】 前記対話エンジンが利用者ごとの対話履歴を記録する対話履歴情報保存部と、対話中の内容に関して最も詳しいと思われる前記第三者を前記対話履歴から選定する介助者選択部とをさらに含み、前記音声情報の内容に対して最適な介助者を選定する請求項 1 から 5 のいずれか一項に記載の音声対話システム。

【請求項 7】 前記音声情報仲介部において前記利用者の対話が円滑に行われていないと判断した場合に、介助者に対してかかる事実を通知する介助依頼通知部をさらに含み、

前記介助依頼通知部において、前記利用者の対話が円滑に行われていないと判断した事実が通知された場合には、前記介助者が前記利用者との対話を自発的に行うことができるとともに、前記介助者と前記利用者との対話において前記介助者のみの音声所定の時間以上継続していることを検知した場合には、前記対話エンジンが前記介助者のみと対話する請求項 1 記載の音声対話システム。

【請求項 8】 前記対話履歴情報保存部に格納されている前記対話履歴を介助者に表示出力する対話履歴表示部と前記介助者からの介助指示を受けつける介助指示部をさらに含み、

前記介助指示部が前記介助者から前記介助指示を受けつけると、前記音声情報仲介部は、前記介助者と前記利用者との間の対話を可能とし、前記介助者と前記利用者との対話において前記介助者の介助の度合が所定のしきい値を超えると、

前記対話エンジンが前記介助者のみと対話する請求項 1 記載の音声対話システム

【請求項 9】 利用者端末から利用者の音声情報を入力する工程と、  
前記音声情報について音声認識を行い、前記音声情報の内容を解析する工程と

前記音声情報の内容に応じて前記音声情報の伝達経路を制御する工程と、  
前記合成音声を出力する工程とを含み、

前記音声情報の伝達経路を制御する工程において、知識データベースを参照して前記音声情報に対応した応答内容を抽出して、前記応答内容に応じた合成音声を作成する音声対話方法であって、

前記音声情報の伝達経路を制御する工程において、前記利用者の対話が円滑に行われているか否かを常時監視し、前記利用者の対話が円滑に行われていないと判断された場合には、前記利用者と前記対話エンジンとの会話への他の端末からの第三者の参加を許可することを特徴とする音声対話方法。

【請求項 10】 利用者の音声情報を入力するステップと、  
前記音声情報について音声認識を行い、前記音声情報の内容を解析するステップと、

前記音声情報の内容に応じて前記音声情報の伝達経路を制御するステップと、  
前記合成音声を出力するステップとを含み、

前記音声情報の伝達経路を制御するステップにおいて、知識データベースを参照して前記音声情報に対応した応答内容を抽出して、前記応答内容に応じた合成音声を作成する音声対話方法を実現するコンピュータに実行させるプログラムを記録したコンピュータ読み取り可能な記録媒体であって、

前記音声情報の伝達経路を制御するステップにおいて、前記利用者の対話が円滑に行われているか否かを常時監視し、前記利用者の対話が円滑に行われていないと判断された場合には、前記利用者と前記対話エンジンとの会話への他の端末からの第三者の参加を許可することを特徴とするコンピュータに実行させるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、利用者がシステムにより生成される合成音声との対話によって種々の機能を実現する音声対話システムを円滑に機能させるための音声対話システム及び音声対話方法に関する。

【0002】

【従来の技術】

昨今のコンピュータ処理速度の向上、扱うことができる情報容量の拡大、並びに音声認識技術の急速な進歩によって、単語レベルの音声認識であれば実用的に問題なく使用できるようになってきたことから、音声認識技術もシステム構築上の重要な要素技術の一つになりつつある。かかる音声認識技術を活用することで、高度なグラフィカルインタフェースを用いるまでもなく、専ら音声を利用した音声対話インタフェースを用いて、種々の課題を解決することが可能となっている。

【0003】

このように音声対話インタフェースを用いることで、利用者が意図する手順によって課題解決を行うことができ、かつ人間が本来有しているオーラルコミュニケーションの能力をフルに活用することができる音声対話システムが、様々な分野において開発されてきている。

【0004】

図1に、従来の音声対話システムの構成図を示す。図1において、1は音声入力部であり、利用者が音声情報を入力するものである。入力媒体としては、マイク等の他、ウェブファイルとしてのファイル渡し等も考えられる。また、入力された音声情報はネットワーク7等を通じて音声認識部2へ送られる。ネットワーク7としては、インターネットを初めとして、WAN/LAN等あらゆる接続形態が考えられる。

【0005】

2は音声認識部であり、入力された音声情報の内容を解析するものである。また、入力された音声にノイズが含まれている場合等に対処するべくノイズリダク



ション機能を有することも考えられる。

【0006】

6は音声出力部であり、音声入力部1からの音声入力に対応した応答を合成音声として出力するものである。なお、出力方法として音声による出力に限定されるものではなく、応答内容を表示装置に表示するものであっても良い。

【0007】

3は音声情報仲介部を示し、音声入力部1及び音声出力部6と対話エンジン4との間の応答タイミングを制御するものである。4は対話エンジンであり、音声入力の内容に応じて知識データベース5を参照することによって、最も適した応答内容を抽出するものである。したがって、対話エンジン4の性能自体が音声対話システムの性能に直接影響を及ぼすことになる。

【0008】

しかしながら、上述したような従来の音声対話システムでは、現在の音声認識技術が100%の認識精度を有するに至っていないことから、利用者の発話意図が十分に認識されない場合も考えられる。また、いかに高性能の対話エンジン4であろうと、人間が日常行っているような自由会話に完全に対応させることは不可能であり、対応しきれない内容の会話になることも十分考えられる。

【0009】

このような場合、当該会話内容に詳しい第三者やオペレータを会話に参加させることで対話エンジン4のみでは適合できなかった会話についても対応することができるようになるものと考えられる。

【0010】

例えば、特公平7-19239号公報においては、音声入力に特定の単語が含まれていることを検知したら第三者であるオペレータと対話できるように切り替えるシステムについて開示している。

【0011】

また、特開平8-76965号公報においては、利用方法が分からなくなったときに、利用者の意思によってオペレータの介在を要求できる音声認識システムについて開示されており、特開平10-124086号公報においては、エキス

パートシステムのみで応答できない場合にシステムサポータによって直接応答できるようにする音声対話システムが開示されている。

【 0 0 1 2 】

【発明が解決しようとする課題】

しかしながら、例えば特公平 7 - 1 9 2 3 9 号公報に開示されている発明では、特定の単語が含まれていなくてもオペレータの介助が必要な場合も想定され、またあらゆる場合を想定した単語を抽出することが實際上困難であることから、実際にかかるシステムを構築することが困難であるという問題点があった。

【 0 0 1 3 】

また、特開平 8 - 7 6 9 6 5 号公報に開示されている発明では、利用者が意図しなければオペレータの介助を受けることができず、第三者が監視していれば無駄な対話であると容易に判断できる対話内容を、相当時間経てからでないとオペレータの介助を受けることができないという問題点があった。

【 0 0 1 4 】

さらに、特開平 1 0 - 1 2 4 0 8 6 号公報に開示されている発明においては、エキスパートシステムにおけるルールデータベースに含まれていない対話の場合にのみシステムサポータの介助を得ることができるが、ルールデータベースに含まれている対話であっても、利用者の知識レベルとルールデータベースの知識レベルが合致していなければ、円滑な対話が成立せず、利用者にとってわかりにくい音声対話システムになってしまうという問題点があった。

【 0 0 1 5 】

本発明は、上述したような問題点を解消するべく、利用者の知識レベルに対応して、オペレータの介在するタイミングを自動的に制御することができる音声対話システム及び音声対話方法を提供することを目的とする。

【 0 0 1 6 】

【課題を解決するための手段】

上記目的を達成するために本発明にかかる音声対話システムは、利用者端末から利用者の音声情報を入力する音声情報入力部と、音声情報について音声認識を行い、音声情報の内容を解析する音声認識部と、音声情報の内容に応じて音声情

報の伝達経路を制御する音声情報仲介部と、知識データベースを参照して音声情報に対応した応答内容を抽出して、応答内容に応じた合成音声を作成する対話エンジンと、合成音声を出力する音声情報出力部とを含む音声対話システムであって、音声情報仲介部は、利用者の対話が円滑に行われているか否かを常時監視し、利用者の対話が円滑に行われていないと判断した場合には、利用者と対話エンジンとの対話への他の端末からの第三者の参加を許可することを特徴とする。

## 【 0 0 1 7 】

かかる構成により、知識データベースの知識レベルと利用者の知識レベルに乖離が存在する場合であっても、オペレータの介助を受けることができるタイミングを自動的に制御することが可能となる。

## 【 0 0 1 8 】

また、本発明にかかる音声対話システムは、音声情報仲介部が、利用者の対話が円滑に行われているか否かを、知識データベースに応答内容が見つからなかった回数の累積値が設定限界回数を越えたか否かによって判断し、設定限界回数を越えた場合には、利用者と対話エンジンとの対話への他の端末からの第三者の参加を許可することが好ましい。一度該当項目が知識データベースに登録されていなかったからといってオペレータの介助が即座に必要というものでもないが、ある程度の回数試行錯誤しても応答として適当な項目にヒットしない場合については、やはりオペレータによる介助が必要となるからである。

## 【 0 0 1 9 】

また、本発明にかかる音声対話システムは、音声情報仲介部において、利用者の対話が円滑に行われているか否かを、対話エンジンが応答してから利用者が反応を返すまでの平均反応時間によって判断し、平均反応時間が第1のしきい値を超えた場合、あるいは第2のしきい値より小さい場合には、利用者と対話エンジンとの対話への他の端末からの第三者の参加を許可することが好ましい。利用者の平均反応時間は、対話が円滑に進行している間は一定の範囲内に収束するものであるが、対話が冗長になってくると次第に反応時間が短くなる傾向にあり、逆に対話の内容についていけなくなると次第に反応時間が長くなる傾向にあるため、かかる対話特有の性質を利用したものである。

## 【 0 0 2 0 】

また、本発明にかかる音声対話システムは、音声情報仲介部において、利用者の対話が円滑に行われているか否かを、利用者の入力した音声情報の平均的な音質の評価値によって判断し、平均的な音質の評価値が第1のしきい値を超えた場合、あるいは第2のしきい値より小さい場合には、利用者と対話エンジンとの対話への他の端末からの第三者の参加を許可することが好ましい。利用者の入力する音声の高さやピッチ等の平均値は、対話が円滑に進行している間は一定の範囲内に収束するものであるが、対話が冗長になってくると次第に高くなる傾向にあり、逆に対話の内容についていけなくなると次第に低くなる傾向にあることから、かかる対話特有の性質を利用したものである。

## 【 0 0 2 1 】

また、本発明にかかる音声対話システムは、音声情報仲介部において、利用者が対話を開始してからの対話時間及び対話エンジンとのアクセス度数に応じて対話進捗度を定め、対話進捗度が低い順に第三者の参加形態が、第三者へ利用者との対話内容を表示して第三者によって対話内容を更新する介入、第三者が利用者と並行して入力する並行入力、第三者が直接利用者と対話する切替と、順次変化することが好ましい。対話進捗度に応じて、利用者が必要とするレベルに応じた介助を提供することができるからである。

## 【 0 0 2 2 】

また、本発明にかかる音声対話システムは、対話エンジンが利用者ごとの対話履歴を記録する対話履歴情報保存部と、対話中の内容に関して最も詳しいと思われる第三者を対話履歴から選定する介助者選択部とをさらに含み、音声情報の内容に対して最適な介助者を選定することが好ましい。対話の内容に応じて、介助専門のオペレータのみならず、対話に参加したことのある利用者の中からも最適な者を選択することができるからである。

## 【 0 0 2 3 】

また、本発明にかかる音声対話システムは、音声情報仲介部において利用者の対話が円滑に行われていないと判断した場合に、介助者に対してかかる事実を通知する介助依頼通知部をさらに含み、介助依頼通知部において、利用者の対話が

円滑に行われていないと判断した事実が通知された場合には、介助者が利用者との対話を自発的に行うことができるとともに、介助者と利用者との対話において介助者のみの音声が入所定の時間以上継続していることを検知した場合には、対話エンジンが前記介助者のみと対話することが好ましい。介助専門のオペレータが自発的に対話を介助することができる機会を付与するとともに、介助専門のオペレータが対話のイニシアティブを取ってしまった事実を検出することで、介助専門のオペレータと直接対話させた方が、より対話が円滑に進行するものと考えられるからである。

## 【 0 0 2 4 】

また、本発明にかかる音声対話システムは、対話履歴情報保存部に格納されている対話履歴を介助者に表示出力する対話履歴表示部と介助者からの介助指示を受けつける介助指示部をさらに含み、介助指示部が介助者から介助指示を受けつけると、音声情報仲介部は、介助者と利用者との間の対話を可能とし、介助者と利用者との対話において介助者の介助の度合が入所定のしきい値を超えると、対話エンジンが介助者のみと対話することが好ましい。同様に、介助専門のオペレータが自発的に対話を介助することができる機会を付与するとともに、介助専門のオペレータが対話のイニシアティブを取ってしまった事実を検出することで、介助専門のオペレータと直接対話させた方が対話が円滑に進行するものと考えられるからである。

## 【 0 0 2 5 】

また、本発明は、上記のような音声対話システムの機能をコンピュータの処理ステップとして実行するソフトウェアを特徴とするものであり、具体的には、利用者端末から利用者の音声情報を入力する工程と、音声情報について音声認識を行い、音声情報の内容を解析する工程と、音声情報の内容に応じて音声情報の伝達経路を制御する工程と、合成音声を入力する工程とを含み、音声情報の伝達経路を制御する工程において、知識データベースを参照して音声情報に対応した応答内容を抽出して、応答内容に応じた合成音声を作成する音声対話方法であって、音声情報の伝達経路を制御する工程において、利用者の対話が円滑に行われているか否かを常時監視し、利用者の対話が円滑に行われていないと判断された場

合には、利用者と対話エンジンとの会話への他の端末からの第三者の参加を許可する音声対話方法並びにそのような工程をプログラムとして記録したコンピュータ読み取り可能な記録媒体であることを特徴とする。

【 0 0 2 6 】

かかる構成により、コンピュータ上へ当該プログラムをロードさせ実行することで、知識データベースの知識レベルと利用者の知識レベルに乖離が存在する場合であっても、オペレータの介助を受けることができるタイミングを自動的に制御することができる音声対話システムを実現することが可能となる。

【 0 0 2 7 】

【発明の実施の形態】

以下、本発明の実施の形態にかかる音声対話システムについて、図面を参照しながら説明する。図 2 は本発明の実施の形態にかかる音声対話システムの構成図である。なお、図 1 と共通する構成要素については同一の符号を付することで、詳細な説明は省略する。

【 0 0 2 8 】

図 2 において、従来と大きく相違するのは、利用者による対話状況を常時監視するように対話管理部 2 1 を設けている点である。対話管理部 2 1 によって、利用者による対話を監視しつつ、対話履歴情報を取得して、対話履歴情報保存部 2 2 へ保存している。

【 0 0 2 9 】

対話履歴情報としては、利用者ごとに音声認識部 2 により認識された対話の内容のみならず、利用者の応答に要する時間や応答までの反応時間、入力される音声の高さやピッチ等の音質に関する情報、また音声情報仲介部 3 を通じて取得できる情報として、知識データベース 5 にヒットしなかった回数の累積値等も含むものとする。

【 0 0 3 0 】

そして、利用者の応答時間からは、利用者に対話を開始してからの累積時間を算出し、入力される音声の高さやピッチ等からは、利用者の入力音声ごとの平均値を算出する。知識データベース 5 にヒットしなかった回数の累積値等を含めた

これらの結果に基づいて、利用者による対話が円滑に行われているか否か、あるいは対話の進捗度がどの程度であるか等を判断することができる。

【0031】

また、音声情報仲介部3においては、新たな機能として、所定の条件を具備しているか否かを判断して、条件を具備している場合には利用者の対話が円滑に行われていないものと判断して、利用者の対話を介助するためのオペレータ端末を接続するオペレータ端末接続部23を起動する機能が付加される。

【0032】

ここで判断すべき所定の条件としては様々な条件が考えられる。まず、利用者の対話が円滑に行われているか否かを、知識データベース5を照会した際に応答内容として適当な項目が登録されていなかった回数（ヒットしなかった回数）の累積値が設定限界回数を超えたか否かによって判断することが考えられる。

【0033】

すなわち、応答内容として適当な項目が登録されていなかった場合には、応答として別の入力を促すような出力がなされることから、何度か試行錯誤するうちに応答として適当な項目にヒットすることもある。したがって、一度該当項目が登録されていなかったからといってオペレータの介助が即座に必要というものでもない。反面、何度試行錯誤しても応答として適当な項目にヒットしない場合については、やはりオペレータによる介助が必要となることから、ヒットしなかった回数を累積して、累積値が所定のしきい値を超えた場合には、利用者の対話にオペレータが参加できるようにオペレータ端末を接続するオペレータ端末接続部23を起動するものである。

【0034】

また、利用者の対話が円滑に行われているか否かを、対話履歴情報保存部22に保存されている対話履歴情報に基づいて算出する利用者の応答ごとの平均反応時間によって判断することも考えられる。

【0035】

すなわち、利用者の反応時間は、対話が円滑に進行している間は一定の範囲内に収束するものであるが、対話が冗長になってくると次第に反応時間が短くなる

傾向にあり、逆に対話の内容についていけなくなると次第に反応時間が長くなる傾向にある。かかる対話特有の性質を利用して、平均反応時間が通常の対話における反応時間の上限である一定のしきい値を超えた場合、あるいは通常の対話における反応時間の下限である一定のしきい値より小さい場合には、利用者の対話にオペレータが参加できるようにオペレータ端末を接続するオペレータ端末接続部 2 3 を起動するものである。

## 【 0 0 3 6 】

さらに、利用者の対話が円滑に行われているか否かを、対話履歴情報保存部 2 2 に保存されている対話履歴情報に基づいて算出する利用者の入力した音声情報の平均的な音の高さやピッチ等によって判断することも考えられる。

## 【 0 0 3 7 】

すなわち、利用者の入力する音声の高さやピッチ等の平均値は、対話が円滑に進行している間は一定の範囲内に収束するものであるが、対話が冗長になると次第に高くなる傾向にあり、逆に対話の内容についていけなくなると次第に低くなる傾向にある。かかる対話特有の性質を利用して、利用者の入力する音声の高さやピッチ等の平均値が通常の対話における音声の高さやピッチ等の平均値の上限である一定のしきい値を超えた場合、あるいは通常の対話における音声の高さやピッチ等の平均値の下限である一定のしきい値より小さい場合には、利用者の対話にオペレータが参加できるようにオペレータ端末を接続するオペレータ端末接続部 2 3 を起動するものである。

## 【 0 0 3 8 】

次に、本発明の実施の形態にかかる音声対話システムを実現するプログラムの処理の流れについて説明する。図 3 に本発明の実施の形態にかかる音声対話システムを実現するプログラムの処理の流れ図を示す。

## 【 0 0 3 9 】

図 3 において、まず利用者が音声を入力すると（ステップ S 3 0 1）、音声認識を行って、入力された音声情報の内容を解析する（ステップ S 3 0 2）。そして、その結果を対話履歴情報として保存する（ステップ S 3 0 3）。同時に、対話開始時以外における平均反応時間、及び入力音声ごとに音声の高さやピッチ等



の平均値を算出して、同じく対話履歴情報として保存する（ステップ S 3 0 4）。

【 0 0 4 0 】

次に、入力音声の内容に基づいて知識データベースを参照して（ステップ S 3 0 5）、入力音声の内容に対応する情報が知識データベースにあれば（ステップ S 3 0 6 : Y e s）、当該情報に基づいて応答内容を自動生成して合成音声として出力する（ステップ S 3 0 7）。このようにして対話を継続することになる。

【 0 0 4 1 】

一方、入力音声の内容に対応する情報が知識データベースになければ（ステップ S 3 0 6 : N o）、再度他の入力を促すメッセージを表示するとともに（ステップ S 3 0 8）、知識データベースをヒットしなかった回数として対話履歴情報中で累積する（ステップ S 3 0 9）。

【 0 0 4 2 】

そして、対話履歴情報において、知識データベースをヒットしなかった回数の累積値が一定のしきい値を超えた場合には（ステップ S 3 1 0 : Y e s）、オペレータによる介助が必要であるものと判断して、オペレータ端末を当該対話へ接続する（ステップ S 3 1 1）。

【 0 0 4 3 】

以上のように本実施の形態によれば、知識データベースにおける知識ベースと利用者の知識レベルとの乖離が大きい場合に、オペレータによる介助を受けることができるタイミングを自動的に制御することが可能となる。

【 0 0 4 4 】

また、対話履歴情報保存部 2 2 に保存されている対話履歴情報に基づいて、利用者が対話を開始してからの対話時間や、対話エンジン 4 にアクセスした回数の累積値を求めることもできる。これら対話時間や対話エンジン 4 にアクセスした回数の累積値が大きな値となっている場合には、利用者の対話の進捗度が高くなっているものと考えられる。

【 0 0 4 5 】

そこで、対話時間や対話エンジン 4 にアクセスした回数の累積値を評価値とし

た対話進捗度に基づいて、オペレータによる利用者への介助方法を変えることも考えられる。すなわち、図4に示すように、対話時間が短い場合や対話エンジン4にアクセスした回数の累積値が小さい場合においては、対話進捗度が低いものと判断し、オペレータ自身によって介助の必要性を判断する方法で十分であると考えられる。

## 【0046】

具体的には、オペレータ端末24に対して利用者に出力する前に一度対話内容を表示し、オペレータが何らかの介助が必要であると判断した場合に限り、出力内容を更新することで十分である。対話エンジン4に対するアクセスをさらに繰り返すことによって、オペレータを介することなく対話を進行できる可能性も高いからである。このようにオペレータによる任意的な介助方法を「介入」と呼ぶ。

## 【0047】

なお、オペレータ端末24における対話内容の表示は対話履歴表示部26によって行われ、オペレータによる介助入力については、音声によって音声入力部25から入力される。

## 【0048】

また、対話時間がある程度累積されてきた場合や対話エンジン4にアクセスした回数の累積値がある程度の値となっている場合においては、対話進捗度がある程度高くなってきたものと判断して、オペレータによるより積極的な介助が必要となっているものと考えられる。

## 【0049】

具体的には、オペレータ端末24に対して対話内容を表示することは同じであるが、利用者にも同時に出力し、オペレータが第三者的立場から新たな出力を入力することも考えられる。ある程度対話が進捗しているにもかかわらず対話が円滑に行われていないと判断されているということは、対話エンジン4に対するアクセスを繰り返すのみでは不十分な箇所が発生しているものと考えられるからである。このようにオペレータによる積極的な介助方法を「並行入力」と呼ぶ。

## 【0050】

なお、この場合も同じく、オペレータ端末 2 4 における対話内容の表示は対話履歴表示部 2 6 によって行われ、オペレータによる介助入力については、音声によって音声入力部 2 5 から入力される。

## 【 0 0 5 1 】

さらに、対話時間が相当累積されている場合や対話エンジン 4 にアクセスした回数の累積値が非常に高くなっている場合においては、対話進捗度がかなり高いものと判断して、オペレータによる積極的な介助が必要となっているものと考えられる。具体的には、オペレータが直接利用者と対話することが考えられる。対話エンジン 4 に対するアクセスでは、対話の続行が不可能になっている可能性が高いものと考えられるからである。このようにオペレータが直接介助するべく対話エンジン 4 を用いない介助方法を「切替」と呼ぶ。

## 【 0 0 5 2 】

なお、この場合も同じく、オペレータ端末 2 4 における対話内容の表示は対話履歴表示部 2 6 によって行われ、オペレータによる介助入力については、音声によって音声入力部 2 5 から入力される。

## 【 0 0 5 3 】

このようにすることで、対話の進捗度に応じて、オペレータによる介助手段に緩急をつけることができ、利用者にとって必要とされるレベルのオペレータによる介助を自動的に判断することが可能となる。

## 【 0 0 5 4 】

また、オペレータは専属のオペレータである必要は必ずしもなく、例えばネットワーク上で同一アプリケーション環境にアクセスしている他の利用者のうち、音声情報の内容に詳しいと考えられる他の利用者であっても良い。この場合、図 5 に示すように、利用者の対話内容履歴が保存されている対話履歴情報保存部 2 2 を照会することで、以前に行われた同一又は類似内容に関する対話における他の利用者のうち、当該内容に最も詳しいと考えられる他の利用者を選定する介助者選択部 5 1 が必要となり、音声情報仲介部 3 が利用者端末の接続先を決定することになる。

## 【 0 0 5 5 】

すなわち、図 5 において、音声情報仲介部 3 が、入力された音声情報の内容に関する過去に行われた対話を対話履歴情報保存部 2 2 を照会することで抽出し、介助者選択部 5 1 において当該対話内容の中から最も当該内容に詳しい者と考えられる他の利用者を選択して、当該他の利用者の端末を接続できるように他の利用者端末接続部 5 2 を起動させる。

## 【 0 0 5 6 】

そして他の利用者端末 5 3 では、それまでの対話履歴が対話履歴表示部 5 5 に表示され、当該他の利用者の知識でもって、音声入力部 5 4 からオペレータの代わりに介助行為を行うことになる。

## 【 0 0 5 7 】

また、対話履歴情報保存部 2 2 に保存されている対話履歴情報に基づいて、様々な方法によって利用者が対話を円滑に進めることができていないことを検出することができるが、この場合にオペレータの自主的な判断によって利用者の対話を介助することも考えられる。

## 【 0 0 5 8 】

すなわち、図 6 に示すように利用者が対話を円滑に進めることができていないことを検出すると同時に、オペレータ端末接続部 6 1 を起動させ、介助依頼通知部 5 6 からオペレータに介助依頼を自動的に通知することによって、オペレータが介助を行えるようにするものである。

## 【 0 0 5 9 】

なお、図 6 に示すように、利用者が対話を円滑に進めることができていないことを検出するまでもなく、常時オペレータ端末接続部 6 1 を起動させておき、オペレータ端末 6 2 における対話履歴表示部 6 4 に利用者と対話エンジンとの間の対話履歴を表示させておくことで、オペレータ自身が介助が必要と判断したタイミングで介助を開始するものであっても良い。

## 【 0 0 6 0 】

この場合、オペレータ端末 6 2 において、さらに介助指示部 6 5 を設け、オペレータ自身の判断で介助が必要であると判断した場合に、オペレータが入力した介助指示を音声情報仲介部 3 の介助指示部 6 5 が受けつけ、対話管理部 2 1 へ渡

すことになる。そして、介助指示を受け取った対話管理部 21 は、オペレータと利用者との間の対話を可能にすることで、オペレータによる介助行為が可能となる。

#### 【0061】

さらに、オペレータによる介助が主体となってしまった場合、例えば利用者と対話エンジンとの対話がほとんどなくなり、オペレータと対話エンジンとの対話のみになってしまったことを検出した場合には、対話エンジンとオペレータの対話のみが行えるように接続を切り替えることも考えられる。すなわち、オペレータによる介助の度合が所定のしきい値を超える程度にまで及んだことを検知できた場合には、利用者端末との接続を切断してオペレータとの対話のみにすることにより、当該問題の早期解決を図るものである。

#### 【0062】

この場合、オペレータによる介助が主体となる状況の判断指標としては、単位時間当たりの情報発信量や、単位時間当たりの無音時間比率、単位時間当たりの話者交代頻度、あるいは単位時間当たりのオペレータ及び利用者の発生時間の比率等の指標が考えられる。なお、特にこれらに限定されるものではなく、オペレータによる介助が主体であると判断できる指標であれば何でも良い。

#### 【0063】

なお、本実施の形態においては、音声字体の入力による対話システムについて説明しているが、同様のシステムがチャットのような文字入力システムにおいても実現可能である。すなわち、チャットにおいては、音声により入力すべき対話の内容がキーボード等によるテキスト入力に変更されているにすぎないものと考えられるからである。この場合、チャットによるやり取りの頻度等を検出する指標としては、例えば単位時間当たりに入力されるテキストデータ量や通信されるテキストデータ量等が考えられるが、基本的には本実施の形態における考え方と大差はない。

#### 【0064】

また、本発明の実施の形態にかかる音声対話システムを実現するプログラムを記憶した記録媒体は、図 7 に示す記録媒体の例に示すように、CD-ROM 72

ー 1 やフロッピーディスク 7 2 - 2 等の可搬型記録媒体 7 2 だけでなく、通信回線の先に備えられた他の記憶装置 7 1 や、コンピュータ 7 3 のハードディスクや R A M 等の記録媒体 7 4 のいずれでも良く、プログラム実行時には、プログラムはローディングされ、主メモリ上で実行される。

【 0 0 6 5 】

また、本発明の実施の形態にかかる音声対話システムにより生成された対話履歴情報等を記録した記録媒体も、図 7 に示す記録媒体の例に示すように、C D - R O M 7 2 - 1 やフロッピーディスク 7 2 - 2 等の可搬型記録媒体 7 2 だけでなく、通信回線の先に備えられた他の記憶装置 7 1 や、コンピュータ 7 3 のハードディスクや R A M 等の記録媒体 7 4 のいずれでも良く、例えば本発明にかかる音声対話システムを利用する際にコンピュータ 7 3 により読み取られる。

【 0 0 6 6 】

【発明の効果】

以上のように本発明にかかる音声対話システムによれば、知識データベースにおける知識ベースと利用者の知識レベルとの乖離が大きい場合においても、オペレータによる介助を受けることができるタイミングを自動的に制御することが可能となる。

【 0 0 6 7 】

また、本発明にかかる音声対話システムによれば、対話の進捗度に応じて、オペレータによる介助手段に緩急をつけることができ、利用者にとって必要とされるレベルのオペレータによる介助を自動的に判断することが可能となる。

【図面の簡単な説明】

【図 1】 従来の音声対話システムの構成図

【図 2】 本発明の実施の形態にかかる音声対話システムの構成図

【図 3】 本発明の実施の形態にかかる音声対話システムにおける処理の流れ

図

【図 4】 介助レベル決定の模式図

【図 5】 本発明の実施の形態にかかる音声対話システムの構成図

【図 6】 本発明の実施の形態にかかる音声対話システムの構成図

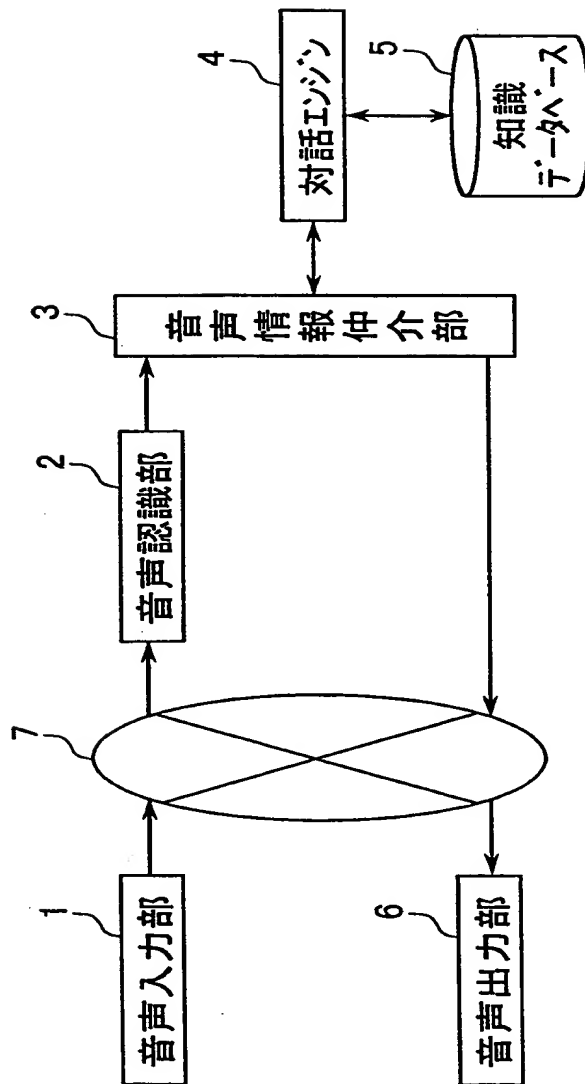
【図 7】 記録媒体の例示図

【符号の説明】

- 1、25、54、63 音声入力部
- 2 音声認識部
- 3 音声情報仲介部
- 4 対話エンジン
- 5 知識データベース
- 6 音声出力部
- 7 ネットワーク
- 21 対話管理部
- 22 対話履歴情報保存部
- 23 オペレータ端末接続部
- 24 オペレータ端末
- 26、55、64 対話履歴表示部
- 51 介助者選択部
- 52 他の利用者端末接続部
- 53 他の利用者端末
- 56 介助依頼通知部
- 61 オペレータ端末接続部
- 62 オペレータ端末
- 65 介助指示部
- 71 回線先の記憶装置
- 72 CD-ROMやフロッピーディスク等の可搬型記録媒体
- 72-1 CD-ROM
- 72-2 フロッピーディスク
- 73 コンピュータ
- 74 コンピュータ上のRAM／ハードディスク等の記録媒体

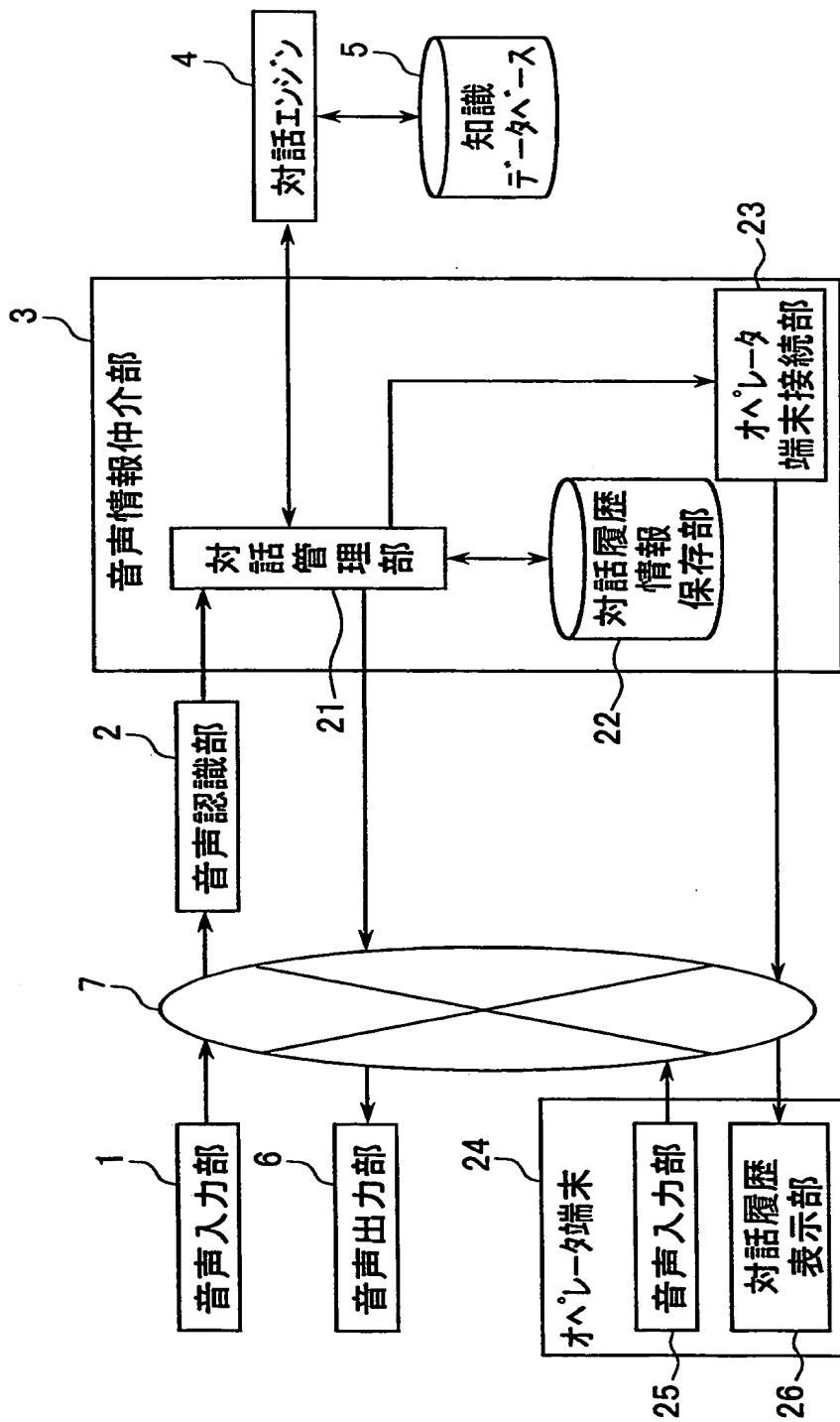
【書類名】 図面

【図 1】

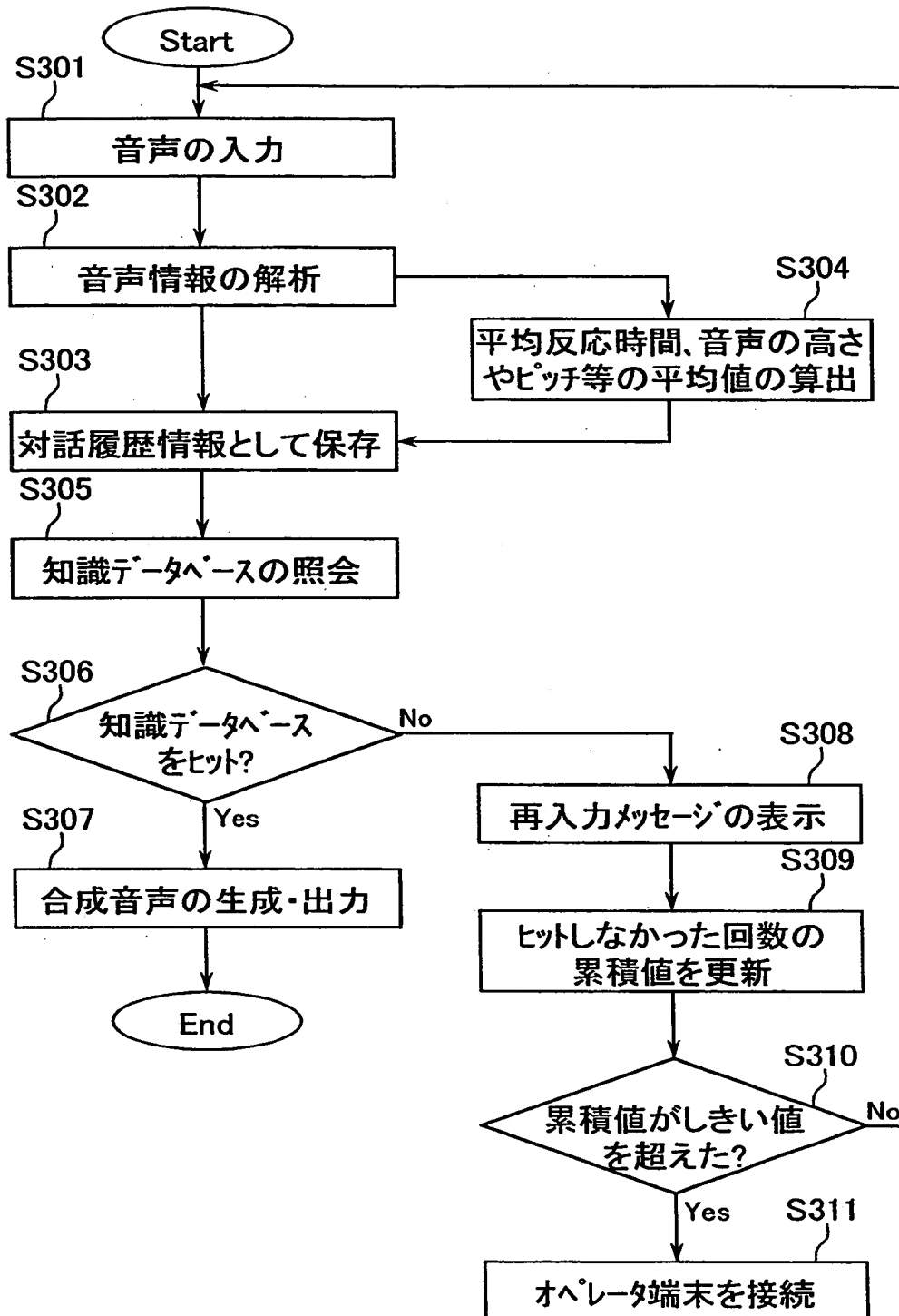




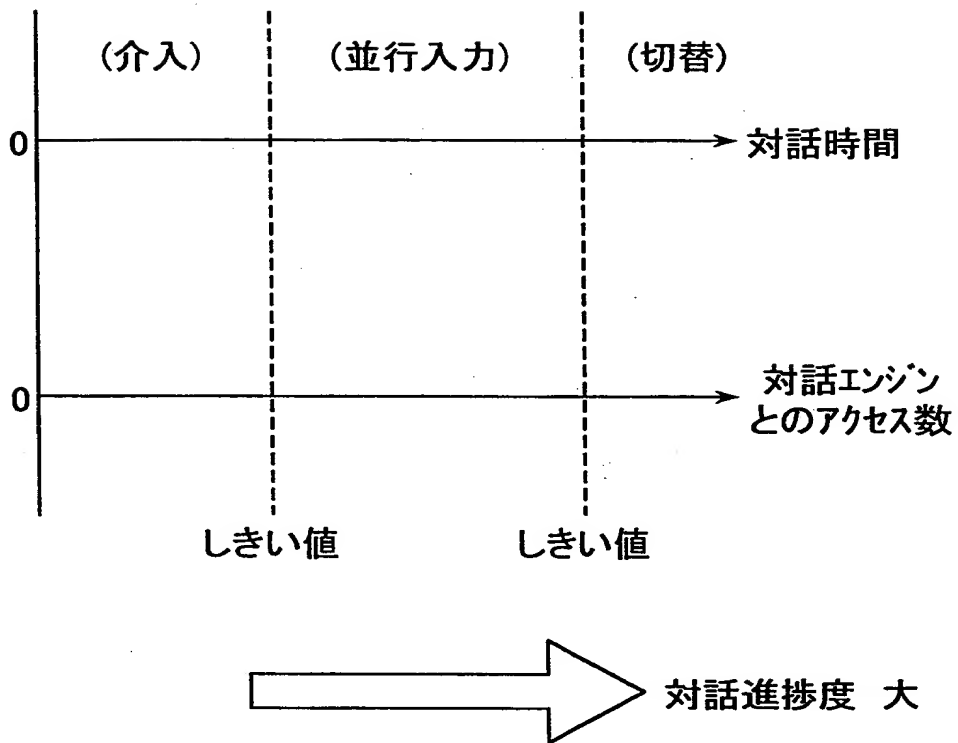
【図2】



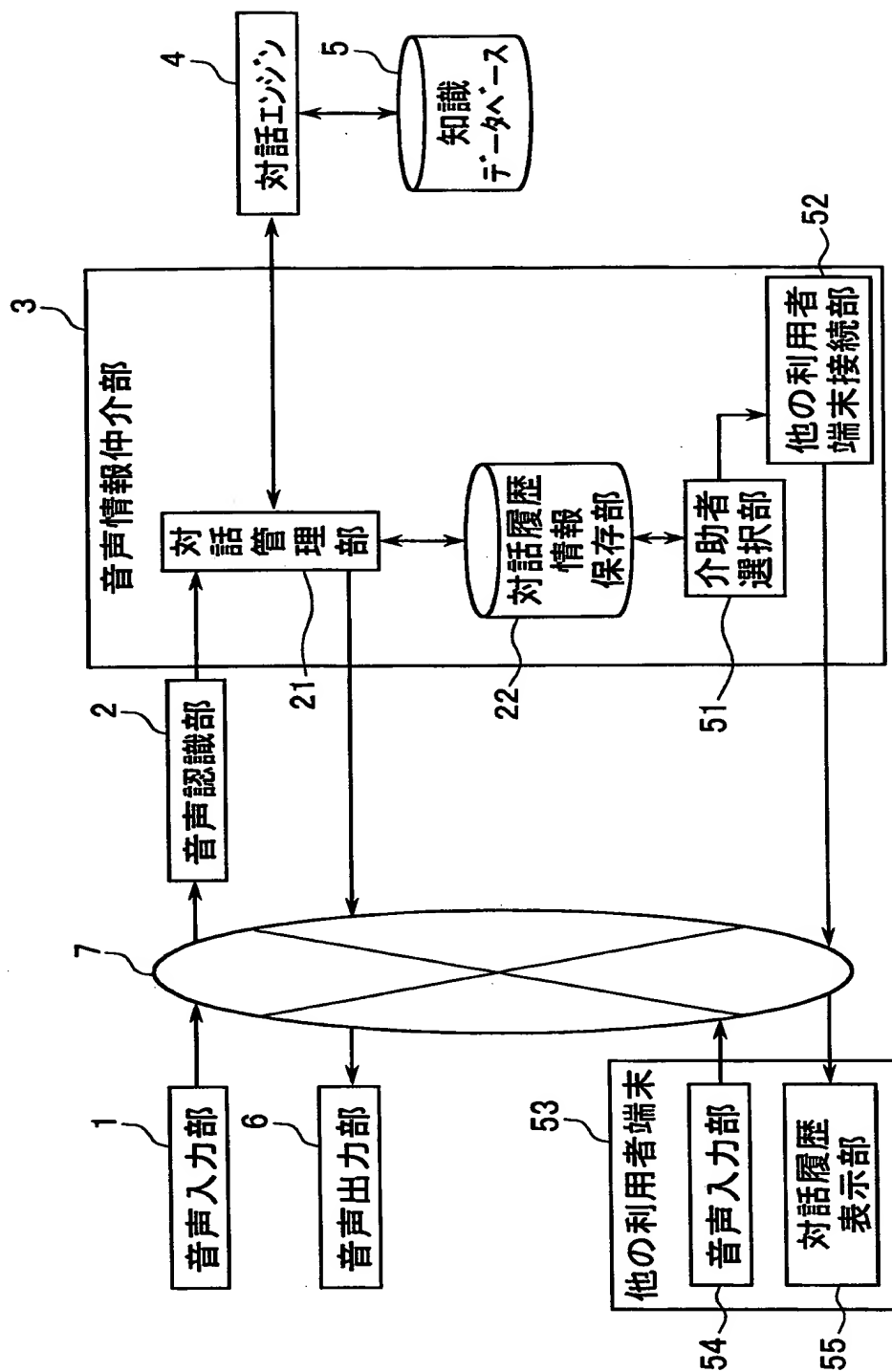
【図 3】



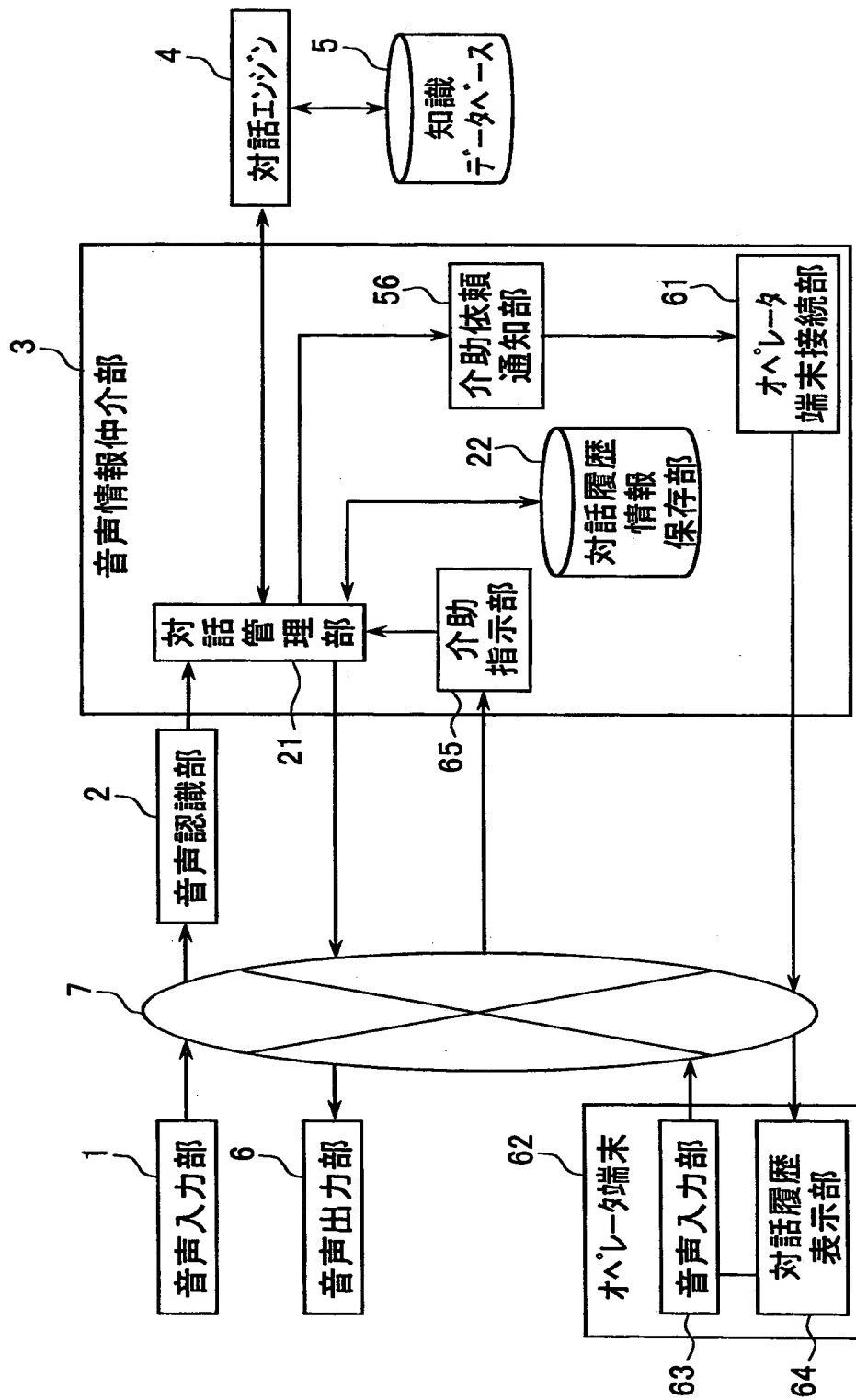
【図 4】



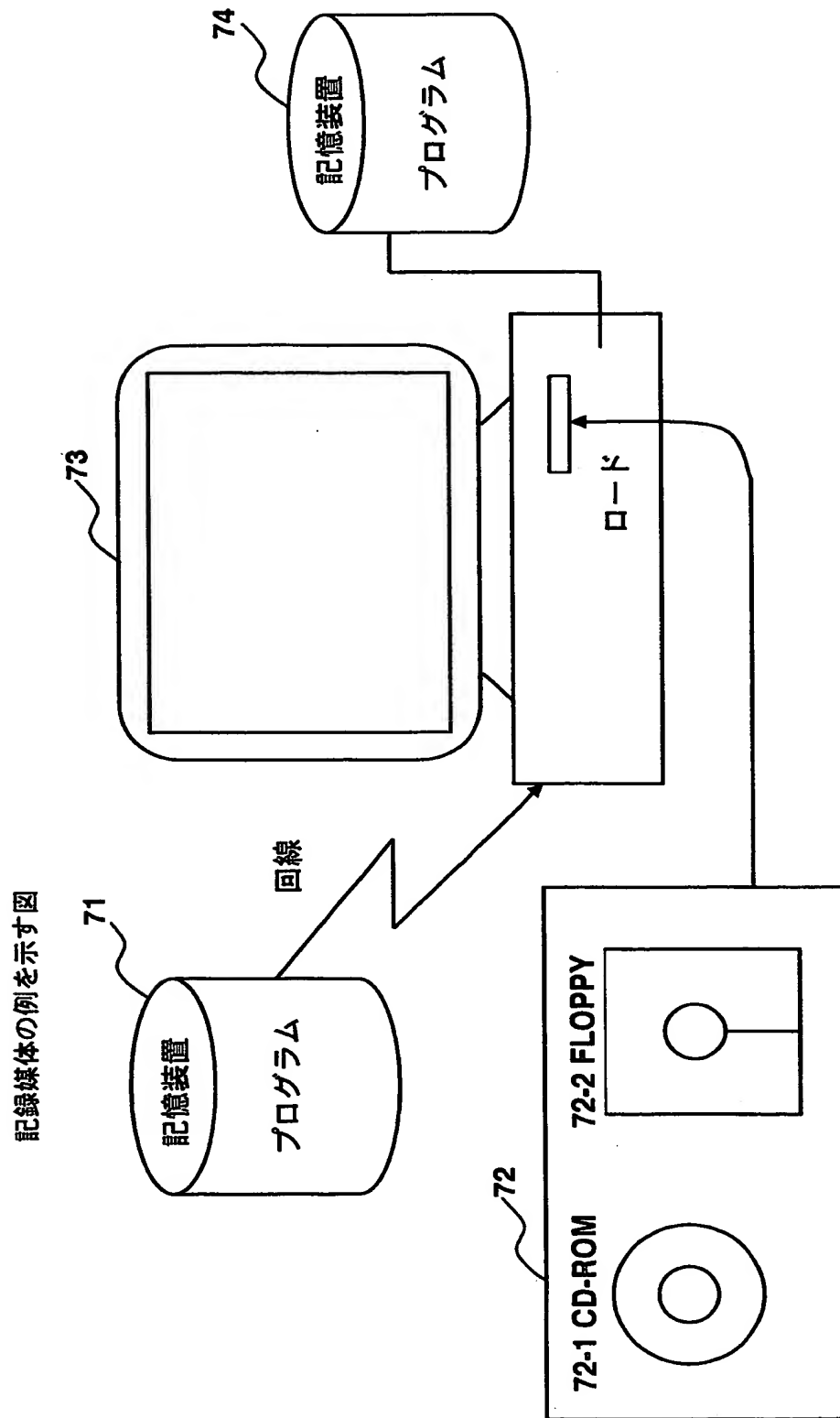
【図5】



【図 6】



【図 7】



【書類名】 要約書

【要約】

【課題】 利用者の知識レベルに対応して、オペレータの介在するタイミングを自動的に制御することができる音声対話システム及び方法を提供する。

【解決手段】 利用者の音声情報を入力し、音声情報について音声認識を行い、音声情報の内容を解析し、音声情報の内容に応じて音声情報の伝達経路を制御し、合成音声を出力するとともに、知識データベースを参照して音声情報に対応した応答内容を抽出して、応答内容に応じた合成音声を作成する音声対話システムであって、利用者の対話が円滑に行われているか否かを常時監視し、利用者の対話が円滑に行われていないと判断された場合に、利用者の会話に第三者の参加を可能にする。

【選択図】 図 2

出 願 人 履 歴 情 報

識別番号                      [ 0 0 0 0 0 5 2 2 3 ]

1. 変更年月日                1 9 9 6 年    3 月 2 6 日

    [変更理由]                住所変更

        住 所                神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号

        氏 名                富士通株式会社